Combined Mobile Robots Motion Control Using Information as Voice and Gesture Commands from Audio and Thermal Visual Sensors

Snezhana Pleshkova[1], Zahari Zahariev[2]

[1, 2] Department of Telecommunications, Technical University, snegpl@tu-sofia.bg

***Abstract*— Motion is one of the main functions of the mobile robots [1]. The problem to achieve the precise control of the mobile robot motion doesn't have a unique solution. This is because there are varieties of situations of mobile robot motion: from motion with full human guidance to mobile robot motion self-control usually tracking some object or human.**

Keywords*— *Mobile Robots, Gesture Commands,Audio and Thermal Visual Sensors

Motion is one of the main functions of the mobile robots [1]. The problem to achieve the precise control of the mobile robot motion doesn't have a unique solution. This is because there are varieties of situations of mobile robot motion: from motion with full human guidance to mobile robot motion self-control usually tracking some object or human. Therefore, it is necessary to have and/or to develop a lot of appropriates methods satisfying the different tasks of mobile robot motion control. One of the wide spread situation of mobile robot motion is to follow the preliminary defined trajectory [2], to avoid obstacles [3] and to go to some object usually marked for example from a human. There are also the tasks when the motion of mobile robot is under human control [4]. In these cases usually the human send the appropriate commands about the way the robot to follow in the motion, the obstacle to avoid, the operations to execute, etc. The sensors on the mobile robot platform receive these commands and the robot must to interpret them. The goal in this article is to propose and develop the human control of mobile robots motion using combined information as voice and gesture commands received from mobile robot's audio and thermal visual sensors.

For voice recognition is most appropriate to use microphone array, so we would be able to receive a full range of incoming voice stream. There are numerous reasons why is better to use microphone array than a single microphone – smaller size for the same sensitivity, we can receive commands from each angle (360^{o} field of listening).

For gesture recognition we can use a lot of different methods: regular VGA camera, IR camera, gesture recognition sensor (Time of flight sensor – ToF [5]) etc. For the purpose of this paper we will use a 32x24 pix IR camera. This will help us to receive information even with low light even without any. This is an advantage against the regular CMOS VGA cameras, and the price is the reason to choose this against ToF sensors [6].

I. The Proposed Block Schema and Algorithm of Combined Motion Control from Human Voice and Gestures to Mobile Robot

On Figure 1 is presented the block schema of the proposed combined mobile robot motion control with voice and gesture commands from a person (human). It is seen from Figure 1 that human can control motion of mobile robot in two ways: sending the voice or gesture commands to the robot. The existed on mobile robot platform audio sensors (on Figure 1 Microphone array) and visual sensors (on Figure 1 Thermo-visual sensor) receive audio and thermo-visual information. The choice of thermal instead of optical visual sensors is based on possibility to extend the application of gesture commands in night situation of mobile robot motion control. The information from audio and thermo-visual sensors is processed in the mobile robot computer system to recognize the voice and gesture commands, needed to achieve human control in each current step of mobile robot motion. The proposed on Figure 1 block schema gives to the human the following possibilities to choose and/or combine the following ways when performing mobile robot motion control: sending only voice commands to the robot; sending only gesture commands to the robot; sending the combination of voice and gesture commands to the robot.

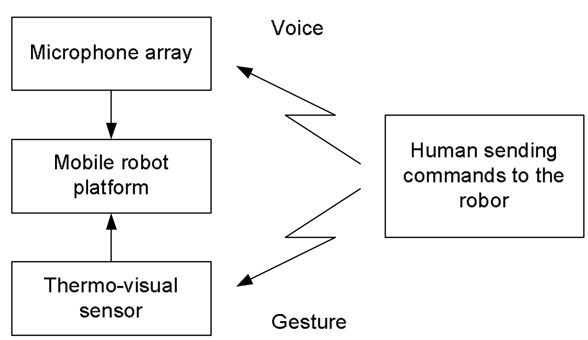

Figure 1. Block schema of combined human mobile robot motion control with voice and gesture commands

The voice control is based on direct comparison between given command and predefined and recorded etalon command. For this purpose system training is needed for each individual operator of the robot.

On Figure 1. Is shown the initialization phase where the operator makes a list of commands by recording them in the robot's memory. These commands are stored at a directory for etalon (preliminary recorded) voice records.

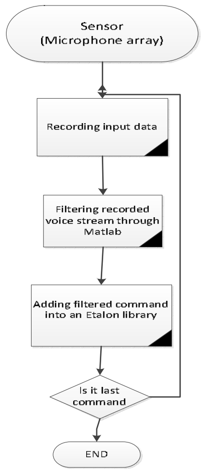

Figure 2. Algorithm for collecting Etalon commands

Gesture recognition algorithm is separated in two different stages – learn stage of algorithm (presented on Fig.3) and test stage of recognition of received gesture commands algorithm (presented on Fig.4).

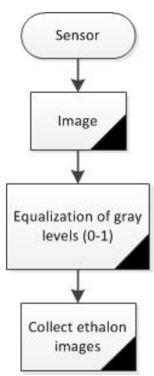

Figure 3. The learn stage of the gesture recognition algorithm.

From IR sensor are taken the IR images of the gestures. After that the captured IR images are sent to computer for brightness levels adjustment using the appropriate grading of pixel intensity from "0" (cold temperature corresponding to background) to "1"(warm temperature corresponding to the example of hand gesture). After that step a new image is stored in a memory of Etalon images.

For the next step is needed to evaluate the similarity of the gestures. This is why it is repeat the etalon gesture several times. Using Matlab implementation of the algorithm shown on Figure 3. it is proposed in the experiments presented in this article to prepare direct comparison of the observed images. For other new input images of the same gesture the difference between Etalon ones can be calculated and used as parameter for decision whether or not the Image difference – "mean(mean(A)) new input images belong to desired gesture. If this parameter is less than some threshold with the values determined experimentally (0.03 in the carried out here experiments and based on real measurements – 10 times repeated the same gesture), therefore the images belong to the same gesture.

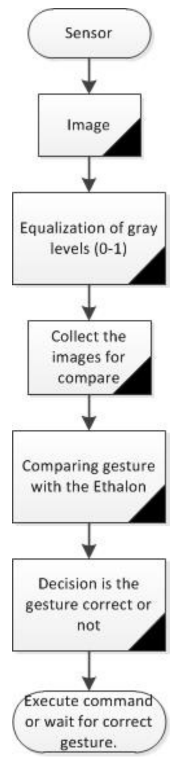

Figure 4. Test stage of the gesture recognition algorithm.

II. EXPERIMENTAL RESULTS

Based on the proposed block schema shown on Figure 1. and on the developed algorithms in Figure 2. are prepared the experiments as software simulations using audio and thermos-visual information from real audio thermo-visual sensors mounted on moving platform of mobile robot.

The experiments are carried out with the following types of devices shown generally as blocks in Fig.1: Mobile robot Surveyor SRV-1 (Figure 10. left) with IR sensor and Wi-Fi module [7]; IR – Sensor MLX90640 (Figure 10. right) with resolution 32x24 pixels [8]; Microphone array with MEMS microphones type STEVAL-MKI155v2 (Figure 9. left) on evaluation board STEVAL-MKI 138v1 (Figure 9. right) [9].

For the purposes of basic control of a mobile robot I will use only few commands, but with the same principle the algorithm can be expanded for more gestures and commands for movement control of the mobile robot:

- Go, Stop, Left, Right

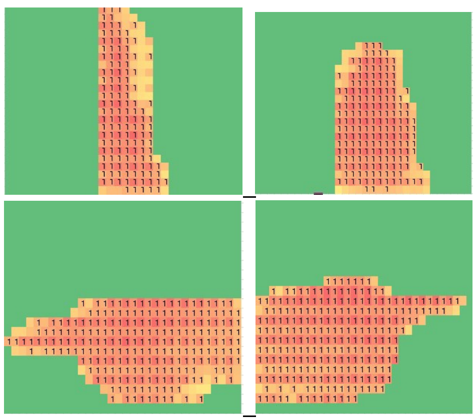

Figure 5. Gestures used for navigating the mobile robot: Go (upper left), Stop (upper right), Left (bottom left), Right (bottom right)

For these gestures the following voice commands looks as follow:

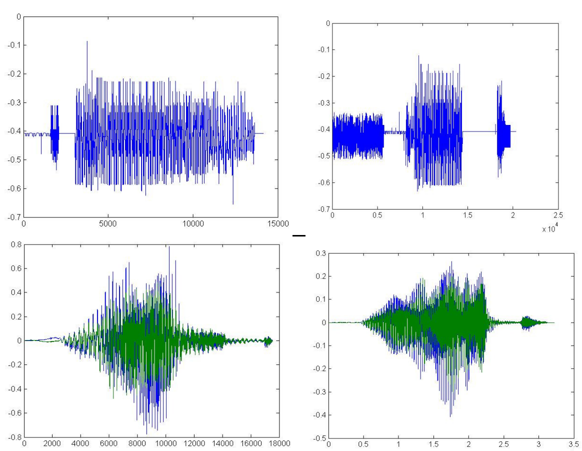

Figure 6. Plot of the voice commands for navigating the mobile robot Amplitude vs time:
Go (upper left), Stop (upper right), Left (bottom left), Right (bottom right)

A. *Voice commands*

For the purpose of the current research I will use a MEMS microphone array – Figure 9. The algorithm shown on Figure 2. is the same as principle as the one for gesture commands. The purpose is to make a library of aurio files, where each command is linked to specific action – movement, stop, awaiting instruction etc. For the purpose is used Matlab Simulink model fo collecting data from the microphone of the mobile robot trough WIFI channel, then this command is stored in memory. After repeating one command for 10 times we can compare these ten record and based on that to set a similarity coefficient. If a given command is similar to some recorded one more than 95% we can say that it is a correct one. Based on that you can see a visual comparisment of command recorded in the library of our robot and a received one on Figure 7.

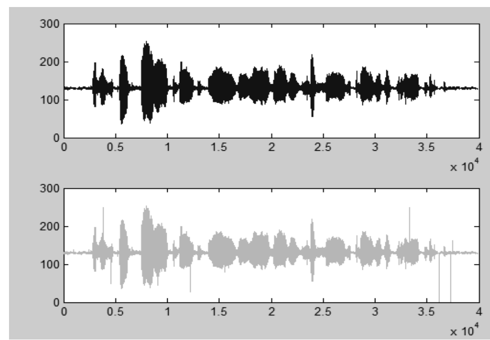

Figure 7. Recorded command (top) vs received one (bottom)

You can see that they are not the same but after we calculate the difference and clear only the mismatches we can evaluate the received audio signal, as you can see on Figure 8.

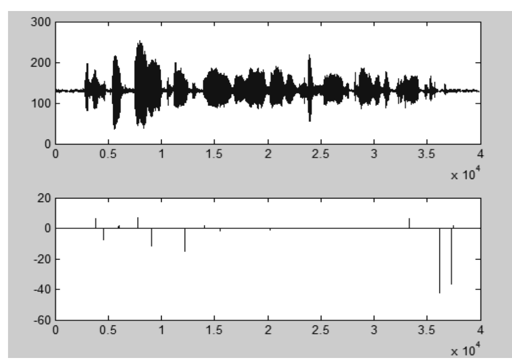

Figure 8. Recorded command (top) vs mismatch of the received audio signal (bottom)

With a simple Matlab compare we see that signals from Figure 5 are 96.3% . According the limit we said we can say that the received audio signal is the same as the etalon one. The robot will execute the connected to this command function and/or action

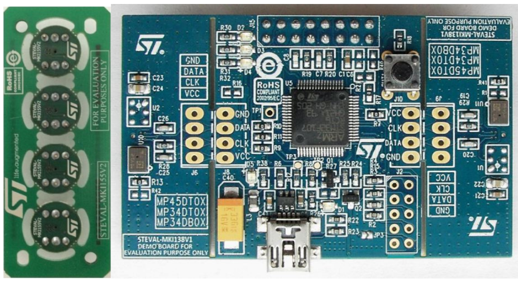

Figure 9. MEMS microphone array STEVAL-MKI155v2 on evaluation board STEVAL-MKI 138v1

B. Gesture commands

The arranged with the above mentioned devices schema for receiving and processing audio and thermo-visual information is tested and used in the experiments for interpretation and processing of audio and visual information in tasks of mobile robot motion control with recognized gesture and voice commands send from human to the robot. The results from these experiments are summarized and described here briefly (only algorithm for the gesture recognition shown in right side in Figure 2.) to demonstrating the proper preliminary processing of audio and thermos-visual information and translating this results as commands to the mobile robot.

Figure 10. Mobile robot Surveyor SRV-1

On Figure 11. is presented as example the received from IR – Sensor MLX90640 thermo-visual image (as numerical data shown in pseudo colors) of a chosen gesture.

Figure 11. Thermo-visual image (as numerical data shown in pseudo colors) chosen as gesture received from IR – MLX90640 thermo-visual sensor.

On Figure 10. and Figure 11. are shown the results from processing thermo-visual information to achieve the correct gesture recognition using the proposed algorithm on Figure 4. implemented in Matlab [10]. All thermo-visual gesture images shown on Figure 12. and Figure 13. are presented after pseudo colors to black and white transformation. The correct gesture recognition is demonstrated on Figure 12. On the left is thermo-visual gesture image used as reference image in the recognition. The middle on Figure 12. is the presented the same (but with little variations) thermo-visual gesture image capture from IR sensor of the mobile robot used as test to correct recognition. On the right is thermo-visual gesture image as difference between two thermo-visual gesture images for which the calculated value of difference is very small - 0.023 and is used as criterion of decide for a correct chosen gesture recognition.

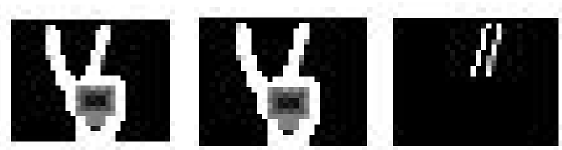

Figure 12. Correct gesture recognition of a chosen thermal image: on the left - reference gesture; on the middle - variation of the same gesture; on the right - image difference between two gestures for which the calculated value of difference is very small - 0.023, i.e. the recognition is correct.

The example of wrong gesture recognition is demonstrated on Figure 13. On the left is shown also the reference thermo-visual gesture image used in the recognition. The middle image on Figure 13. is also the same (but with larges variations) image of gesture capture from IR sensor of the mobile robot used as test to wrong recognition. On the right is image difference between two gestures for which the calculated value of difference is bigger - 0.13 and is used as criterion for decision of wrong gesture recognition.

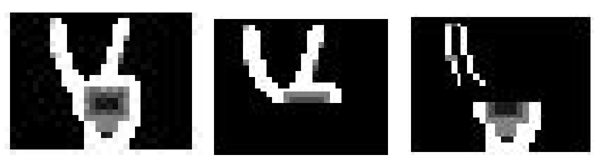

Figure 13. Wrong gesture recognition of a chosen thermal image recognition: on the left - reference gesture; on the middle - variation of the gesture; on the right - image difference between two gestures for which the calculated value of difference is bigger - 0.13, i.e. the recognition is wrong.

III. CONCLUSION

The following conclusions can be listed based on the achieved result in the tests and experiments:

- the theoretical bases of using the probabilities of errors in voice and gesture commands recognition and the existence of correlation dependence between these commands, leads practically to flexibility and precision in realization of mobile robot motion control from human;

- the usefulness and the advantage of sending combined voice and gesture commands to the robot is the increase the precision and effectiveness of motion control from human;

- it is achieved an effective human-robot interface: human interact with the robot deciding to send only voice or only gesture or combined voice gesture commands, depending on human observation whether or not the mobile robot execute correctly current motion step recognizing correct each of these sending to the robot command.

- improvements in the proposed algorithm can be in terms of training with various and numerous variations of gestures and voice commands, and also to investigate the influence of the joint usage of gestures and voice commands in enhancement the accuracy of mobile robot motion control from human commands.

ACKNOLEDGMENTS

This work was supported by Technical University – Sofia inner program to research projects under 172PD0035-07:" Development of methods and algorithms for recognition of gestures and speech, and processing to interpret the received commands".

REFERENCES

[1] P. Soueres, and J. D. Boissonnat, "Robot motion planning and control", New York, Springer, 1998.

[2] W. Xu, J. Wei, J. M. Dolan, H. Zhao, and H. Zha, "A real-time motion planner with trajectory optimization for autonomous vehicles", In IEEE International Conference on Robotics and Automation (ICRA), , 2012, pp. 2061 – 2067.

[3] S. Dadelahi, M. R. J. Motlagh, and A. R. M. Shahri, "Obstacle Displacement Prediction for Robot Motion Planning and Velocity Changes", International Journal of Information and Electronics Engineering, Vol. 3, No. 3, May 2013, pp. 270 - 274.

[4] T. S. Jin, J. M. Lee, and H. Hashimoto, "Position Control of Mobile Robot for Human-Following in Intelligent Space with Distributed Sensors", International Journal of Control, Automation, and Systems, vol. 4, no. 2, April 2006, pp. 204-216.

[5] https://en.wikipedia.org/wiki/Time_of_flight.

[6] https://www.melexis.com/en/product/MLX75023/Automotive-QVGA-Time-of-Flight-Sensor